

# Towards a Poetics of Strangeness: Experiments in Classifying Language of Technological Novelty

Michael Simeone, Advait Gundavajhala Venkata Koundinya, Anandh Ravi Kumar and Ed Finn

09.08.17

*Peer-Reviewed By: Ted Underwood*

*Clusters: Genre*

*Article DOI: 10.22148/16.015*

*Dataverse DOI: 10.7910/DVN/MSXKNB*

The trajectory of science fiction since World War II has been defined by its relationship with technoscientific imaginaries.<sup>1</sup> In the Golden Age of the 1930s and 1940s, writers like Isaac Asimov and Robert Heinlein dreamed of the robots and rocket ships that would preoccupy thousands of engineers a few decades later. In 1980s cyberpunk, Vernor Vinge, William Gibson, and Bruce Sterling imagined virtual worlds that informed generations of technology entrepreneurs. When Margaret Atwood was asked what draws her to dystopian visions of the future, she responded, "I read the newspaper."<sup>2</sup> This is not just a reiteration of the truism that science fiction is always about the present as well as the future. In fact, we will argue, science fiction is a genre defined by its special relationship with what we might term "scientific reality," or the set of paradigms, aspirations, and discourses associated with technoscientific research.

This is a feedback loop: science fiction informs technical research even as it draws on that research for themes, fundamental questions, and horizons of possibility. The media of exchange include particular technical ideas (Jules Verne's submarine; H. G. Wells' atomic weapons) as well as aesthetics, characters, and ideologies. There is an appealing strangeness to science fiction that makes the genre a fertile source of inspiration and motivation.

But what is the relationship between strangeness and science fiction? What makes a narrative moment of a science fiction text (or any text, for that matter) feel strange and visionary?

This essay attempts to describe strangeness, a key genre activity of science fiction (SF), with low-level features of language by reading alongside a supervised classification routine. Through this exploration, we hope to gain a new perspective on the broader terrain of the technoscientific imaginary in literature—not by relegating interpretation to machines, but by deliberate application of a model based on a modest subset of SF texts. Accordingly, we see our results as suggestive of SF mechanics, not characteristic of all SF literary production. Through this glimpse of SF mechanics we may gain a deeper understanding of the qualities of science fiction, but it may be more accurate to describe this work as a study of SF sentences, rather than SF works. We believe our results inform a model of strangeness more than they provide a description of this literary genre. However, we also hope to accelerate future work on the discovery and retrieval of key points of SF activity in archives of texts, and, perhaps, to increase the frequency and efficacy of transfer from literary imagination to technoscientific practice.

<sup>1</sup>Work by the Authors was partially supported by The National Aeronautics and Space Administration through the funding of a Workshop on Understanding Literature and Art Cultures for Transformative Research held at Arizona State University in 2014 (funded as a subaward from NASA grant number NNX12AJ32G). The authors also acknowledge the contribution of the workshop participants in framing this research question and Elizabeth Garbee and Jacqueline Hettel for their earlier experiments in language and novelty. Simeone and Finn contributed roughly equal effort to this paper and are its primary authors. Koundinya and Kumar participated as secondary authors through their contributions to the areas of workflow design, data analysis, and implementation of key prototypes.

<sup>2</sup>Brooke Hunter, "How The Handmaid's Tale Taught Me to Imagine Many Possible Futures," Slate, July 29, 2016.

## Background

One of the most compelling and enduring definitions of science fiction was formulated by Darko Suvin in 1972: a literature of cognitive estrangement "whose main formal device is an imaginative framework alternative to the author's empirical environment."<sup>3</sup> Suvin's work harnesses the central tension of science fiction for a productive critical goal, "transmuting an apparently hopeless contradiction into constitutive antinomy," as Gerry Canavan puts it.<sup>4</sup> Suvin defines the genre's essential role at the intersection of competing imaginaries: the space of scientific reality and prediction, on the one hand, and the artistic mode of escaping our shared reality for other worlds and possibilities, on the other. The mechanics of this antinomy depend on the "novum," the particular world-building event or technological object that establishes the difference between a science fiction narrative and accepted reality.

The novum may be understood best not as a literary device but as an analogy to the process of empirical scientific discovery. The reader's experience of the novum mimics that sense of wonder and surprise that accompanies one's first experience with a sociotechnical innovation, followed immediately by the same kind of intellectual reordering of the world and the broader impacts and potential consequences of the new thing that begin to unfurl in thought and in social action. In this light, the novum is a stand-in for the scientific breakthrough or discovery. As early as 1922, William F. Ogburn and Dorothy Thomas speculated that the persistence of multiple discovery—two or more people making the same intellectual leap at roughly the same time, like Isaac Newton and Gottfried Wilhelm von Leibniz with calculus—signal that there is an important sociological component to technoscientific innovation.<sup>5</sup> Robert K. Merton advanced this theory of "multiples," and it continues to attract attention in the guise of "evolutionary epistemology" (applying Darwinian mechanics to the spread of ideas) and Richard Dawkins's "memetics."<sup>6</sup> As Robert Heinlein put it in *The Door Into Summer*, "When railroading time comes you can railroad—but not before."<sup>7</sup>

This, at least, is the technoscientific romance of the novum. After all, Suvin adapts the novum from Ernst Bloch and the Marxist tradition of utopia, situating science fiction's "essentially epistemological function [as] the imagination of alternative social and economic forms," as Fredric Jameson argues.<sup>8</sup> Suvin diagrams the genre's tension between the pragmatic imaginaries of technologists, engineers, and researchers attempting to expand the boundary lines of technoscientific knowledge and the deliberately new, different, unfamiliar territory that is accessible only when one abandons the "imagination in a straightjacket" of scientific thinking.<sup>9</sup> The novum is a signpost for the genre turn, a signal to the reader of the narrative rule-set but also, crucially, the rule-set for technoscientific innovation. At its apogee, the novum encapsulates the entire science fiction enterprise within a single narrative idea, an iconic or pervasive referent that functions as a technoscientific catalyst: Asimov's robots and Three Laws of Robotics, for example, have informed generations of technical research.<sup>10</sup>

And yet Suvin's framing of the novum encodes a crucial counterpoint to the seemingly discrete status of this literary object, this new thing. He argues that estrangement is a cognitive process, depending on the subject position of a reader, the status of the author, reader, and text in relation to perceived realities that may or may not be shared. Literary cognition draws in language, historical context, subject identity, and their interrelationships over time, from authorship and publication to reception, but also extending forwards and backwards to encompass the full sweep of the narrative in play. As China Miéville puts it in his critique of Suvin, we need to understand cognition "as something done with language by someone to someone."<sup>11</sup> For Miéville, cognition is a damning reintroduction of "capitalist science's bullshit" into the promisingly utopian alterities of the genre because it is utterly contingent on the historical and social context of literary production.<sup>12</sup>

Cognitive estrangement is contingent, is the product of the same linguistic, technoscientific processes that create Popular Mechanics and the Strategic Defense Initiative. Cognitive estrangement evolves over time, drawing resonances from historical experience, technical knowledge, and intersubjective identity to project science fiction imaginaries. It is a contingency of a special kind. Samuel R. Delaney, Robert Markley, Steven Shaviro, and Kim Stanley Robinson have all discussed

<sup>3</sup>Darko Suvin, "On the Poetics of the Science Fiction Genre," *College English* 34, no. 3 (December 1972): 373.

<sup>4</sup>Gerry Canavan, "Defined by a Hollow: Essays on Utopia, Science Fiction and Political Epistemology," Darko Suvin, Oxford: Peter Lang, 2010," *Historical Materialism* 21, no. 1 (January 1, 2013): 210, doi:10.1163/1569206X-12341280.

<sup>5</sup>William F. Ogburn and Dorothy Thomas, "Are Inventions Inevitable? A Note on Social Evolution," *Political Science Quarterly* 37, no. 1 (March 1, 1922): 83–98, doi:10.2307/2142320.

<sup>6</sup>Richard Dawkins, *The Selfish Gene*, 3rd ed (New York: Oxford University Press, 2006).

<sup>7</sup>Robert A. Heinlein, *The Door into Summer* (New York: Doubleday, 1957), 120.

<sup>8</sup>Fredric Jameson, *Archaeologies of the Future: The Desire Called Utopia and Other Science Fictions* (New York: Verso, 2005), xiv.

<sup>9</sup>Finn and Kathryn Cramer, eds., *Hieroglyph: Stories and Visions for a Better Future* (New York: William Morrow, 2014), vii.

<sup>10</sup>Finn and Kathryn Cramer, *Hieroglyph*, xxiii–xvi.

<sup>11</sup>Mark Bould and China Miéville, eds., *Red Planets: Marxism and Science Fiction* (Middletown, CT: Wesleyan University Press, 2009), 235.

<sup>12</sup>Bould and Miéville, *Red Planets*, 240.

science fiction in this way as a form of simulation: a genre that does not represent the human experience but extrapolates it according to a genre rule-set.<sup>13</sup> Shaviro posits this epistemological shift as a literature that "precedes its object: it doesn't imitate or stand in for a given thing, but provides a program for generating it."<sup>14</sup> That modelling function, Shaviro argues, does not require a subject position or a sentient being to play the role of author or observer, just as quantum mechanics does not require a sentient observer in order for a wave function to collapse into determination: any machine will do.<sup>15</sup>

When we consider science fiction as a form of simulation grounded in technical methods of data analysis and extrapolation, we are extending Gwyneth Jones' suggestion that science fiction requires the "appearance of command over the language of science"—quite literally, the careful juxtaposition of specialist vocabulary to achieve the effect of cognitive estrangement.<sup>16</sup> According to this logic, a good science fiction novel manages to unite literary and information-theoretical notions of surprise: the narrative revelation of new, probabilistically unlikely information. This approach to simulation and the instrumental role of language offers a clear pathway to the methodology we detail below: can we trace these juxtapositions and create a set of rules for identifying them procedurally?

We have identified Support Vector Machines (SVMs) as a method that could help illuminate and describe these juxtapositions. Linear classifiers like our SVM approach have an established use in text classification at the document level, as they handle effectively the great number of features offered by natural language data.<sup>17</sup> There has been work exploring how to use classifiers like SVMs to compare sentences derived from the same documents, but containing different stylistic tendencies.<sup>18</sup> Additionally, SVMs demonstrate an ability to identify nuanced distinctions among texts; they have been used to identify expertise level of authors in film review essays,<sup>19</sup> and to identify sentences with very specific semantic registers, such as thanking, questioning, specifying, and suggesting.<sup>20</sup>

## The Strangegram

We began the project by asking ourselves, can we quantify the novum? Is there a salient pattern that describes the way writers create the novelty effect that others have used to define the genre of science fiction? While SF scholars can readily identify nova in texts during hands-on reading and examination, until now there has been no way to operationalize the identification of key moments of invention in documents already identified as science fiction. To describe parts of the text that excite speculation, and discover any common features connecting them, would enrich our understanding of what many critics, in various terms, identify as the core function of science fiction. While we do not intend to contain the idea of the novum within the descriptive capabilities of machine learning, we do wish to explore consistencies that span multiple instantiations of the genre.

Finding commonalities, even within subgenres, serves both the literary and instrumental purposes we outline above. We also do not imagine that strangegrams work independently from historical, generic, and textual contexts. Activated by the genre relationship with the reader, we see these features add up to produce an effect consonant with the novum posited by Suvin and others when theorizing the genre of science fiction.

We initially hypothesized that there are enough consistencies across estranging moments of SF texts to support automated identification of SF nova. Specifically, these consistencies would manifest in a way that might be closely related to Suvin's own formulation, namely through word occurrence (specific words associated with strangeness) and co-occurrence (words

<sup>13</sup>Robert Markley, *Dying Planet: Mars in Science and the Imagination* (Durham: Duke University Press, 2005), 356.

<sup>14</sup>Steven Shaviro, *Doom Patrols: A Theoretical Fiction about Postmodernism*, (New York: Serpent's Tail, 1997), 17.

<sup>15</sup>Shaviro, *Doom Patrols*, 122-34.

<sup>16</sup>Gwyneth A. Jones, *Deconstructing the Starships: Science, Fiction and Reality* (Liverpool: Liverpool University Press, 1999), 16.

<sup>17</sup>Thorsten Joachims, "Text Categorization with Support Vector Machines: Learning with Many Relevant Features," in *Machine Learning: ECML-98: 10th European Conference on Machine Learning Chemnitz, Germany, April 21-23, 1998 Proceedings*, ed. Claire Nédellec and Céline Rouveirol (Berlin, Heidelberg: Springer Berlin Heidelberg, 1998), 137-42, <http://dx.doi.org/10.1007/BFb0026683>.

<sup>18</sup>Nitin Jindal and Bing Liu, "Identifying Comparative Sentences in Text Documents." In *Proceedings of the Twenty-Ninth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, vol. 2006, 244-51).

<sup>19</sup>M Jiang, and J Dienser, "Says Who...? Identification of Critic versus Layman Reviews of Documentary Films," (paper presented at the 26th International Conference on Computational Linguistics (COLING), Osaka, Japan, December 11-16, 2016).

<sup>20</sup>Anthony Khoo, Yuval Marom, and David Albrecht, "Experiments with Sentence Classification." In *Proceedings of the 2006 Australasian Language Technology Workshop (ALTW2006, Association of Computational Linguistics, 2006)*, 18-25.

appearing in the same parts of the text or in an unusual combination) in sentences to comprise a kind of detectable, lexical-literary object. Thus, based on the proposed contours of SF genres and subgenres, we posited the "strangegram" as a heuristic for describing and retrieving SF nova.

We contend that there is something strange about the strangegram that is not merely symbolic or cultural. Taking cues from Suvin, we are interested in the combination of specific kinds of words, and how those mixtures cooperate in the production of strangeness. Take for instance Nicola Griffith's "slate" technology from her novel *Slow River*, published in 1995: "Exactly. A slate stuffed with information."<sup>21</sup> In 2016, the idea of a flat, screen-based computer is not "strange" in the strictest sense; these technologies exist and widely circulate. That does not mean that the sentence itself does not maintain an element of strangeness. The combination of words "slate," "stuffed," and "information" not only produces a semantic strangeness for a reader consuming the text before the advent of the iPad, but statistically, this combination of words stands out from the other language around it, as well as from other language in the genre. These are technology words, or words with some denotative association with technoscience, in atypical combinations with words that are not. It is not that tablet computers have become less strange in 20 years. Grasped in terms of the strangegram, it is our own everyday language that continues to bear out this strangeness.

In our first attempts at testing the strangegrams hypothesis, we explored nova as identifiable by strong semantic relationships among a preliminary, constrained vocabulary of technology and what we imagined to be technology-adjacent terms. We used a corpus very similar to the one used for our later experiment (see Data). After our initial -n-gram lists returned nothing notable, we attempted a collocate analysis of words commonly associated with some of the nova we found, and paired that approach with a semantic network analysis that attempted to find similar pairs or triples of concepts present across the examples.

For both collocation and semantic network analyses, we searched for term co-occurrences with a window size of 5L and 5R (looking five words to either side of the word under analysis), and while we did not include stopped words in our results, we did count them in our window size. For instance, even though we stopped out the word "because" in our examples, the word "because" still counted toward the value of 5L or 5R when computing distances among terms.

Our first results were largely inconclusive. We found that while we were able to observe common techniques in certain subgenres (e.g. the combination of two common words into a novel invention such as "rocket wrench" in 1960s pulp SF about space travel), looking for semantic associations among a fixed set of terms did not sufficiently describe, at the language level, a majority of the nova we sampled.

The core elements of science fiction posited by Suvin, it seemed, were not evoked solely in specific combinations and sequences of words to make the constituent parts newly weird. The classic example of "the door irised open," while providing an example of a strangegram, is no blueprint for nova or strangeness. This initial foray into describing nova showed us that strangeness, if it is consistent at all, occurs in a more diffuse pattern and relies on more various narrative techniques than word recombination alone.

Subsequently, we expanded our hypothesis about what a strangegram could be. Rather than consider the fabric of SF nova to be specific kinds of technology and non-technology words falling in close proximity to one another, or words in unusual combination, we hypothesized the strangegram as a critical mass of certain words at the sentence level.

Thus, our most recent version of the strangegram concept contains relaxed expectations about word combination. Namely, specific words do not need to occur together (as in Ngrams), nor do specific kinds of words need to be associated with other kinds of words (as in a kind of collocate analysis). Instead, it is the density of certain words or characters in a moment of the text that help create the strangeness of nova. The current force of the term "strangegram" invokes the "n-gram" designation used in computational linguistics, but with a crucial shift in the dimension of measurement. Based on a probabilistic language model, the approach to identifying an n-gram looks at groups of words that appear together in a contiguous chain and thus behave as a discrete term (e.g. "birthday cake," "problem child," "car accident"). But the genre effect of SF is not reducible to a single term or even a predictable sequence of words. Rather than measuring relationships between words *n* units apart, the strangegram measures a dimension of cognitive estrangement that is more elastically bounded by what we might term a narrative moment. In this analysis, we use the sentence as the boundary of that moment. Thus, we revise our hypothesis to suppose that nova are produced in a field of meaning among certain words, and that field produces an effect of cognitive estrangement.

As a final clarification, our aim is not to comprehensively describe the underlying mechanics inherent in all strangeness

<sup>21</sup>Nicola Griffith, *Slow River*, 1st ed (New York: Del Rey, 1995), 12.

found in SF genres and subgenres; we are sensitive to the potential and real variations in SF across time periods, cultures, and subgenres. We are trying to trace any common symptoms across various subgenres and temporalities, as manifest in low-level features of the text (i.e., the distribution of words over sentences), which may help us better understand what has felt elusive about the novelty effect of SF in the first place. In short, we are looking to describe some level of consistent expression of SF nova amid constellations of historical and cultural variation.

## Data

For our study, we created a training data set from a population of positive and negative cases of technological nova in SF texts. With each example being a sentence, we drew at random positive and negative examples from a subset of science fiction novels published online by Project Gutenberg and the Technovelgy online glossary of science fiction inventions. We aimed to keep about a 50/50 split between positive and negative cases, adjusting for length of sentences. Both positive and negative cases were represented by similar amounts of total words.

Similar to content analysis approaches in social science, we documented our designation of strange sentences to help with consistency. So, when labeling our training and test sentences as "strange" or "not strange," we used the document in Appendix A: Qualitative Description of Strangeness as a guide for use by two scholars of science fiction (authors Finn and Simeone).<sup>22</sup> In an attempt to gauge the robustness of our definition of strangeness, we measured agreement between the two coders and calculated Cohen's Kappa along with total percentage of agreement out of a small, random subset of sentences ( $n = 138$ ). In this assessment, we reached an agreement of 83 percent and a kappa of 0.63. While there is no universal standard for an acceptable kappa score, ours falls in the range of "substantial agreement" according to the general guidelines offered by Landis and Koch.<sup>23</sup>

Given these results, we were satisfied with the overall consistency produced by our definition of strangeness, and for the purposes of our experiment, we used one rater's labels for the full training and test set (Simeone's). However, we see this combination of social science research methods with machine learning to be a promising direction for classification of narrative moments, and we plan on future work that integrates intercoder reliability ratings into training and test evaluations of classifiers.<sup>24</sup>

Crucially, this approach has limitations. Endemic to our (or any) supervised approach is that we can only reproduce the subjective classifications made by the sentence coders. While we cannot make universalizing claims about strangeness, this data can help elucidate that it is possible to define something key to the science fiction at a conceptual level and see that definition bear out when tested with more low-level features of sentences.

## Acquisition and Transformation

We used a 2012 version of the Project Gutenberg Science Fiction Bookshelf, wherein we performed deduplication of stories and novels and eliminated any pieces not written in English.

After removing front and end matter from the Project Gutenberg text files, we cleaned the text to remove errant symbols and lowercased all words. We finished by dividing each novel or short story into sentences, each of which served as the basic unit of analysis.

Similarly, we acquired entries from the site Technovelgy by scripting a download routine that copied each entry into a database. We performed the same operations as above to remove any symbols and standardize letter casing. We also split entries into individual sentences rather than consider whole paragraphs.

When composing the training corpus we ranked all the words in the training set of sentences according to term frequency-indirect document frequency (TF-IDF) and eliminated low-scoring words from the training set. TF-IDF is a statistical

<sup>22</sup>doi:10.7910/DVN/MSXKNB

<sup>23</sup>Richard J. Landis and Gary G. Koch, "The measurement of observer agreement for categorical data," *Biometrics* (1977): 159-174.

<sup>24</sup>For a thorough description of the problem and possible approaches, see Grimmer, Justin, Gary King, and Chiara Superti, *The Unreliability of Measures of Intercoder Reliability, and What to do About it*, working paper, 2015.

measure of the value or relevance of an individual word in a particular document, controlled against the frequency of that word in the full corpus of documents. We used this index to rank all words in a training set and choose only the top 100 ranked terms in each set of strange and not-strange training sentences, before moving to the next stage of analysis. This dramatically reduced the total words available to the classifier when evaluating positive and negative sentences. We also eliminated proper nouns and lowercased all terms.

## Methodology

We attempt to separate language from SF texts into two categories: sentences that create a sense of strangeness and sentences that do not. A key pair of assumptions of our study is that 1) there are some latent, low-level similarities among individual cases of SF nova sampled from different times and places, and 2) those similarities are consistent enough to make predictions and classifications about sentences based on those features. As a result, we explored several approaches to identifying possible commonalities among SF nova, particularly corpus linguistics and text-to-network analysis. We eventually decided on a supervised machine learning routine, which we outline below. We believe these methodological choices model a mode of inquiry at the intersection of traditional humanistic approaches to criticism and the now well-established digital humanities approaches to large-scale corpora analysis. Our study does not ground far-reaching truth claims on its empirical analysis. Rather, we think of this work as producing a modest, suggestive set of results that we hope to use as illustrative, rather than definite evidence of our critical claims.

Much like scholars such as David Bamman and Ted Underwood are able to demonstrate that the structure of certain subsets of literary meaning (such as genre and character) can be understood and predicted as distributions of a select vocabulary of words over a collection of documents,<sup>25</sup> our study uses top words or keywords as features for analysis. While our machine learning approach differs from the Bayesian method used by Bamman and Underwood, we too depend on the core assumption that high-level literary meaning (character, genre, speculation) can be traced consistently across texts by, essentially, counting which words tend to co-appear with one another. As a zoomed-in version of this approach, our examination of the defining characteristics of SF nova explores the distribution of words over sentences, in order to give us insight into the moments in the text that prompt the reader to engage according to the expectations of the SF genre. While in this study we have no capacity to predict all instances of SF nova across all texts through the strangegram, we do aim to show that the strangegram is a unit that makes SF nova more tractable for both critic and machine.

Crucially, the unit of analysis for our study is sentences. We considered over approximately 10,000 total sentences in our training and testing, although the majority of our comparisons dealt with about 1500 sentences apiece. We drew our sentences from novels and short stories provided by Project Gutenberg, as well as descriptions of technology across hundreds of SF works from the Technovelgy collection.

Furthermore, we narrowed our consideration of strangeness to technological novelties. We established a description of the code "S" to assign to sentences for the purpose of training and testing our classifier. The full description of "S" or strange can be found in <img alt="diamond symbol" data-bbox="225 665 238 678"/>

We acknowledge that this does not account for all literary strangeness, but if Suvin and others are to be believed, technology is a major participant in the kind of strangeness important to the activity of SF. Again, our aim here is not to detect all science fiction, but to see what attempts at detection illuminate for us.

## Classification

To explore this revised hypothesis, we used a SVM machine learning algorithm to perform a discriminative classification of sentences present in SF texts. In opposition to generative models such as Latent Dirichlet Analysis, discriminative models perform classification based on present features in the data, rather than a probabilistic model of possible features to be found in both observed and as-yet-unobserved examples. The classification SVM performs is only useful to distinguish one group from another in a given data set. SVM cannot predict, so to speak, what the features of another strangegram might be. Our

<sup>25</sup>David Bamman, Ted Underwood, and Noah A. Smith, "A Bayesian Mixed Effects Model of Literary Character," in 52nd Annual Meeting of the Association for Computational Linguistics (ACL2014 - Proceedings of the Conference, vol. 1, 2014), 370–79.

discriminative model emphasizes observed differences between only the novum and non-novum sentences we supply. For this reason, the discriminative SVM model suits our purposes by producing an up-or-down indication of strangeness that is consistent, at the word level, with confirmed positive or negative cases. While generative and discriminative models are both used in document classification, we are specifically interested in this specific comparison without making generalizations about nova in other SF documents or subgenres. Nor do we proceed as if our training set of nova is the only way, or the best way, to select or even define nova in SF. We ran our classification only to see if a successful distinction could be made between novum and non-novum sentences in our pilot data set. Inferences about the viability of their predictive or descriptive power we leave in the hands of a literary interpretation.

We trained our SVM classifier using the training data set that comprised sentences and Technovelgy entries, minus stopped words and words with relatively low TF-IDF scores. We then trained and tested the classifier using a several combinations of Gutenberg and Technovelgy "strange" and "not strange" sentences. Because each experiment contained approximately 1500 sentences total, we used 10-fold cross-validation to evaluate performance. The overall performance of the classifier can be found in Table 1.

## Results

Our approaches and manual inspections previous to use of SVM showed one of the most obvious approaches to creating a novum is to combine two or more existing technoscientific keywords. Readers of Carey Rockwell's *Treachery in Outer Space* would already be familiar with the terms "rocket" and "wrench," but combining them provides a textbook example of cognitive estrangement on a linguistic level. This works especially well with strangegrams that already have a certain cachet, a world-building power about them, as "rocket" does. The Golden Age was like a nursery for the science fiction linguistics of "rocket," making it a prefix for transforming all sorts of other nouns: rocket man, rocket platform, rocket fuel, rocket tube. One might imagine a caricature of science fiction as a massive OuLiPo machine for generating new combinations of technoscientific jargon and evaluating their success based on the genre marketplace. The description could as easily be applied to William S. Burroughs and the cut-up method as a deliberate process for exposing new and unexpected ideas.

But our subsequently deployed SVM classifier did not consider word order; it worked primarily with vocabulary to make assessments and predictions. This approach showed modest success in classifying strange and non-strange sentences from both collections of source texts.

	Technovelgy NS (747)	Gutenberg NS (840)
Technovelgy S (684)	0.713 (.583)	0.762 (.622)
Gutenberg S (999)	0.896 (.881)	0.837 (.883)

Table 1. F1 Scores for SVM Classifier.

The table above presents the weighted average of F1 scores from comparisons of sentences classified as "strange" and "not strange." Each collection of sentences is followed by the number of instances that comprised our samples. F1 scores are derived by considering the number of correct positive classifications for "strange" over the number of classifications returned as "strange" divided by the actual number of sentences labeled as strange by human coders.

$$F1 = 2 \frac{pr}{p+r} \text{ where } p = \frac{tp}{tp+fp}, r = \frac{tp}{tp+fn}$$

TP = True positive

FP = False Positive

FN = False Negative

We also ran all four experiments by adding a stoplist to our preprocessing workflow. This eliminated common articles, verbs, and adverbs. The performance of the classifier after this additional step is presented in parentheses after the first set of values. Using both stopped and non-stopped text data, we ran a total of eight experiments.

Additionally, when evaluating how the SVM classifier made distinctions between S and NS sentences, we calculated the importance of individual words to the classifier’s decisions. Below is a table of ranked terms/features used for classification. The terms are ranked by information gain per training and test activity. Here we display the top 15 terms for each, and shade the numbers in green according to their values relative to all four experiments. The calculation of information gain uses the method outlined by Fayyad and Irani.<sup>26</sup>

Ts x Tns		Ts x Gns		Gs x Tns		Gs x Gns	
0.08661	the	0.02441	-	0.19396	the	0.10305	ship
0.04691	a	0.01922	its	0.10741	and	0.07567	space
0.03875	of	0.018	had	0.09938	ship	0.06917	the
0.02511	and	0.01441	she	0.09066	of	0.04371	she
0.01515	its	0.01381	don’t	0.0837	space	0.03386	ships
0.01392	in	0.01373	a	0.06137	to	0.02643	of
0.01349	as	0.01257	the	0.05481	was	0.02207	earth
0.01195	to	0.01232	was	0.04483	in	0.02072	her
0.01187	which	0.01169	robot	0.03788	said	0.02024	air
0.01092	by	0.01169	thats	0.03457	their	0.0198	screen
0.01085	from	0.01105	are	0.03272	as	0.01887	car
0.00993	an	0.00846	didn	0.03248	had	0.01868	miles
0.00919	into	0.00806	him	0.02912	ships	0.0174	in
0.00911	-	0.00793	coffee	0.02814	on	0.01682	planet
0.00814	be	0.00791	i	0.02623	were	0.01382	you
0.00701	machine	0.00765	machine	0.02533	earth	0.0138	power
0.00699	is	0.00747	her	0.02502	planet	0.01352	rocket
0.00682	some	0.0073	he	0.023	they	0.01162	sun
0.00658	or	0.00697	these	0.02251	power	0.0116	said
0.00601	on	0.00638	back	0.02111	screen	0.01084	control

Table 2. Information Gain Scores of Features.

## Discussion

Size and sampling are potential confounds for this work. The size of each experiment (approximately 1500 instances) is not large enough to make definitive claims about SF writ large. Furthermore, the collection acquired from Project Gutenberg contains more texts from the 1950s and 1960s than the earlier and later parts of the 20th century. We also limited our definition of “strange” to technology. All of these constraints limit what we can say about our results, but we still observe some meaningful patterns in our preliminary work and the subsequent experiments we ran using an SVM classifier.

From our limited view into the genre, we start to see possible descriptions of what makes SF signal to readers that something is new, and to indicate points in the text that require participatory speculation in order for their genre experience to inhere. The results highlight tendencies about the literary mechanics of the strangegram and the procedural rules by which the crafting of new technologies in science fiction often proceeds.

The words that the classifier found most useful in distinguishing strangegrams from non-strangegrams offer compelling potential insights into the role of syntax in constructing the estrangement of science fiction.

First, many of the words at the top of the above lists fall into the category of function words, or words that play a more grammatical role in the sentence and tend to have ambiguous meaning. The most significant word for the classification in three out of four cases was the definite article “the,” and in the fourth comparison (Technovelgy S vs. Gutenberg NS) it was still a relatively heavily weighted word in 8th place. The definite article is a noun signal, a syntactical setup for some

<sup>26</sup>Usama M. Fayyad, Keki B. Irani, “Multi-interval discretization of continuous valued attributes for classification learning,” in 13th International Joint Conference on Artificial Intelligence (1993): 1022-1027.



kind of object, entity, or idea. The stylistic choice to use the definite article emphasizes the noun in question, grammatically distinguishing it from the characters in the narrative (whose purposes and effects are conveyed through possessive articles).

Another significant cluster of words, prepositions such as "of", "to", "in", "by," etc., furthers this emphasis on objects in relationship to one another. These words structure connections between nouns, operating as both visual and conceptual mechanisms for establishing order, hierarchy, scale, and other comparative relationships in the imagination of the reader. Curious about the apparent importance of some of these words to our model, we ran the same classifiers with a standard stopword list to see how the removal of words like "the" would change the effectiveness of the classifiers. The results are shown in parentheses in Table 1. In three out of four cases, the stopwords are so important that their removal significantly degrades the performance of the classifier.

This result is similar to findings by researchers investigating the quantitative features of literary genres or styles at the word level, including recent work by Sarah Allison, Ted Underwood, and Matt Jockers.<sup>27</sup> Given this body of work, it is not surprising that these kinds of words matter in distinguishing stylistic categories. However, each of these aforementioned studies deal in whole works of literature, whereas ours addressed individual sentences. We are intrigued that these words mattered in drawing a distinction even at the sentence level among sentences, and this finding may indicate a specific narrative style for the production of strangeness within science fiction texts. These words do not define science fiction style, but they do designate narrative moments of strangeness—and we see these effects as importantly distinct.

We believe the prominence of function words further supports our observation that the strangeagram operates not just at the level of meaning, but also of syntax. Future work will consider other syntactical features of the text, but the importance of impersonal description through articles is significant here.

Even in these preliminary results, syntax clearly matters, especially with managing nouns that explicitly denote technological objects. One of the most powerful functions of the genre is to create and manipulate ambiguity at the linguistic and social levels. Not every word in strangeagrams is recognizable as technological in all registers. The word "screen," for example, appears frequently in our results table, playing a wide range of roles. The word is familiar, but context makes it clear that these authors are using it in novel ways, extending, amplifying, and reframing received definitions in order to achieve cognitive estrangement.

Second, hyphens appear to matter. Perhaps a further sign of this kind of ordering, the presence of the hyphen is a major distinguishing feature (#1 in the Technovelgy S vs. Gutenberg NS comparison) in the Technovelgy corpus. It is hard to imagine a more straightforward syntactical vehicle for achieving the strangeagram effect: take two nouns and create a neologism by yoking them together with a hyphen: rocket-ship, ray-gun. But we also note that the hyphen seems to have played no role in the Gutenberg comparisons, perhaps because the stylistic conventions of that corpus and its pulp origins were less amenable to this particular form of word-building.

Third, our results point to the importance of neologisms, which are often recombinant, ambiguous extensions of existing concepts. The "-." character ranked highly in two experiments (both involving the Technovelgy collection), indicating that combining words together is a prevailing strategy for narrating technological strangeness. Importantly, our TF-IDF rankings and limits on total features meant that words like "whipcord," "viewscreen," and "heatgun" were not even considered in classification, even though they too play on the modes of strangeness described above. By considering only hyphenated neologisms, the classifier may miss some of the SF craft in the text, but absent a method for parsing portmanteau words, we believe accounting for hyphenated words helps indicate neologisms and their work in the text without overfitting around a handful of anomalous features. Both words in the term "neck-phone" exist, for example, but the whole scientific apparatus of "phone" succeeds the body part "neck," blending the connotative meanings together in order to invite the reader to simulate a future where "phone" is as multifarious and fungible as the word "rocket" is today. We also observed dashes to signal asides in speech, as well as non-neologistic complex modifiers, both of which serve the purpose of extending existing concepts or ambiguating technology terms to appear plausible in concert with reader imaginations.

Finally, there are examples of words that by themselves do not metonymize an elaborate technological apparatus, as "rocket" and "wrench" surely do. But words like "control," "sun," or "power" (all of which are top-ranked terms for the classifier working with Gutenberg instances of SF), when they appear in the same sentence as other similar words, can produce a strangeagram, or the estrangement of familiar nouns and verbs. For instance, our classifier returned the entry "A second or two later the tube burned out" (from Blake Savage's story "Rip Foster Rides the Gray Planet"). There are no words

<sup>27</sup>See Allison et al., *Quantitative Formalism* (Stanford Literary Lab, 2011); Ted Underwood, *The Life Cycles of Genres* (Cultural Analytics, 2016); Matthew Jockers, *Macroanalysis: Digital Methods and Literary History* (Urbana: University of Illinois Press, 2013), 106-110.

in this sentence that reference a specific device or tool, nor do they invoke any capability that a reader could imagine a technology pathway to reach. Yet the sentence is strange because tubes of many kinds do not typically burn or burn out, and burning in "A second or two" invites speculation about fuel and materials. No single "gram" in this line sufficiently accounts for the strangeness of the sentence. These kinds of results returned by the classifier and validated by SF readers indicate that strangegrams rely on a kind of semantic discord accumulated across many words at a time. Apart from the direct concoction of novel concepts and devices or calling things "strange" outright, unlikely combinations of words in the telling of events produces a strangeness that invites, but does not settle, the speculative faculties of readers.

There are places where the classifier does not account for what human readings uncover to be science-fiction. We also observe in our data set that science fiction creates strangeness through the old-fashioned method of simply describing it as strange. "They went down a narrow room filled with bulky metal objects of bright scarlet or violet that gleamed weirdly..."; "Could the strange clothing be the tie by which the aliens held to him" (Norton, *Time Traders*; emphases added). We see the world through the eyes of a character who shares our state of astonishment and wonder. However, words like "weird" and "strange" were not part of our collection of features (they were ranked out by our TF-IDF preprocessing). The role of these words could very well be the subject of future refinements to our classifiers.

## **Towards a Poetics of Strangeness**

These results demonstrate that (technological) strangeness is, at least for this corpus, a fairly consistent signal, and that its mechanism operates through syntax as much as diction. We see a consistent set of patterns in word deployment, a series of operations that, within our samples of science fiction, the genre executes over and over again. The data suggest that the presence and the ground rules of strangegrams are observable and possibly predictable: explicit grammatical articulation of the novel, recombination, indefinite description, extension-ambiguation of existing technology lexicon, neologisms, explicit uses of the words like "strange" and "weird," and estrangement of familiar nouns and verbs abound in the results of our classification experiment. This consistency suggests that some of the poetics of SF technology are machine-readable. Furthermore, that strangeness can be quantified, even experimentally, opens up similar work with other genres that can explore key feelings or narrative moments. One wonders what a "murdergram" would look like for detective fiction, for instance.

The extent to which this framing of strangeness is contiguous with broader SF as a genre is a subject for future work. To what extent does the sentence-level construction of strangeness depend on the conventions or genre environment of science fiction to flourish, and to what extent do these occurrences of technological strangeness signal dialog with broader technoscientific imaginaries? At the very least, these results hint at broader consistencies for the genre as a form of simulation. Science fiction's genre rules, particularly in so-called "hard science fiction," take on a particular relationship with technical imaginaries. Stories in this genre strategically mimic the literary tics of particular kinds of technical literature—the narrated discoveries, innovations, and observations of scientists and engineers—while reinvesting those discourses with the human concerns, the characters, the sex, and violence that will capture public interest and sell books. Readers of a Kim Stanley Robinson novel expect to get up close and personal with crucial technical ideas and the scientific principles behind them.

At its core, this kind of science fiction takes on the same instrumental role as the science it inhabits: words written to spur particular kinds of imagination at the boundaries of human knowledge. It functions as a vanguard to technical exploration, asking the same "what if?" questions that launch innovative research, but operating in the limitless laboratory of the literary imagination. This is the poetics of strangeness at work, offering up new worlds and ideas through the artful manipulation of core stylistic mechanisms from technoscientific discourse like definite articles, prepositions, and hyphenations.

If we take that thesis to its logical conclusion, it offers two stories for how the simulations of strangeness function in SF. The first is the narrative of cognitive estrangement as it has been interpreted and debated by Suvin, Jameson, and the rest. Strangegrams provide a way to conceive of SF nova as the linguistic equivalent of visceral experience, produced by modifying and remixing the building blocks of how a language typically tells stories. Instead of merely using familiar words to describe the new, or neatly cordoning off new terms or concepts in neologisms, science fiction allows the familiar and the strange to contaminate one another through recognizable patterns. Viewed through the perspective of strangegrams, SF is the genre of those experiences delivered through calculated, articulated strangeness. Bolted together by periodic invocations of the strange, the simulation plays out in our minds as we imagine these alternate worlds and contemplate the technological and ideological conditions of our own existence that make them impossible.

The second mode of simulation takes us back to that “capitalist science’s bullshit” that Miéville rails against. In this story, strangeness is everywhere, not just in SF. It is a sentence-level process, a low-level method for language experimentation that is closely coupled with the technoscientific discourses that generate and evaluate many of the terminological building blocks of the strangegram. The output of this version of strangeness is not only imaginative conceptions of the utopian impossible, but also radical extensions of the possible. Perhaps the output of the science fiction simulation is a broader technoscientific imaginary informed by both the scientific method and Star Trek, with these inflections of the language of technical innovation preceding and at times shaping the direction of technoscientific change.

If this is true, then the traces of the feedback loop between science and science fiction should be visible through the traffic of strangegrams between technical literature, journalism, and science fiction. The much-remarked upon anecdotal evidence is clear: the X Prize Foundation would not have created a Tricorder X-Prize or a Science Fiction Advisory Council without the precursors of Roddenberry and Star Trek; Leo Szilard would not have urged Einstein into instigating the Manhattan Project without H. G. Wells and his visions of atomic war; Amazon engineers would not have developed the Kindle (initially codenamed “Fiona”) without Neal Stephenson’s *The Diamond Age*. But the contours of a larger system of circulation, perhaps even a model for that circulation, remain out of reach—the objective of future research, we hope.

Our experiments with the strangegram suggest a process at work that can be contoured in more concrete descriptions of literary activity than the now-traditional terms “cognition,” “strangeness,” and/or “imagination.” Literary critics understand the political, economic, cultural, historical, and stylistic dimensions of the genre’s imaginative mechanics. But in our research, we extend the impulse to understand SF’s genre mechanisms by experimenting with machine learning: if the word-level workings of SF are traceable at the sentence level, we gain additional insight into how the genre can overlay reality and simulation to such powerful political and imaginative effect. The fact that the novum is traceable by algorithm (at least in the constrained framework of this study) signals the fact that it is a semantic container with low-level features, a set of straightforward verbal and linguistic operations.

What are the consequences of that simplicity for the broader political and literary debates surrounding science fiction? To begin with, we have some evidence that it is easy to mint the fundamental unit of exchange, the utopian semantic currency of strangeness. That facility also lends weight to the notion that this currency circulates beyond the SF where we found it, into the real-world marketplace of ideas. These different literatures have their own styles, their own ideological and semantic contingencies, but the strangegram suggests that they may all rely on the same technology of imagination.

The strangegram is useful not as an intervention in five decades of critical debate about what literary science fiction is for, but rather for further investigating the mechanics of what we have always assumed SF does—illuminate the realities of the present with the spotlight of alterity—at a scale that is granular, rhetorical, and traceable.

Imaginations about technology and culture abound outside and across the borders of what we often designate as SF. We invest so much critical energy debating the alterity of SF that it becomes easy to overlook how much the genre shares with documents and collections considered to be non-literary. One could imagine a number of document types available to strangegram analysis: technoscientific journalism, technical research, policy documents, and other ancillary genres. As both an object for machine learning algorithms and a linguistic currency of imagination, the strangegram provides a poetics of strangeness: how texts create technoscientific novelty with a craft indexed by distributions of words both humble and surprising. Turning toward the very fabric of strangeness can help us understand the continuities and feedbacks among the science fictions we accept to be possible, and the science fictions we accept, every day, to be true.

